Measurement for Outcome and Economic Analysis

Robert M. Kaplan
University of California, San Diego
NIH, OBSSR RCT Course
April 28, 2003

Where This is Going

- Profile Approach
  - SF-36
- Utility Approach
  - QWB
- Preference Assessment
- Cost/Effectiveness Analysis

Question 1

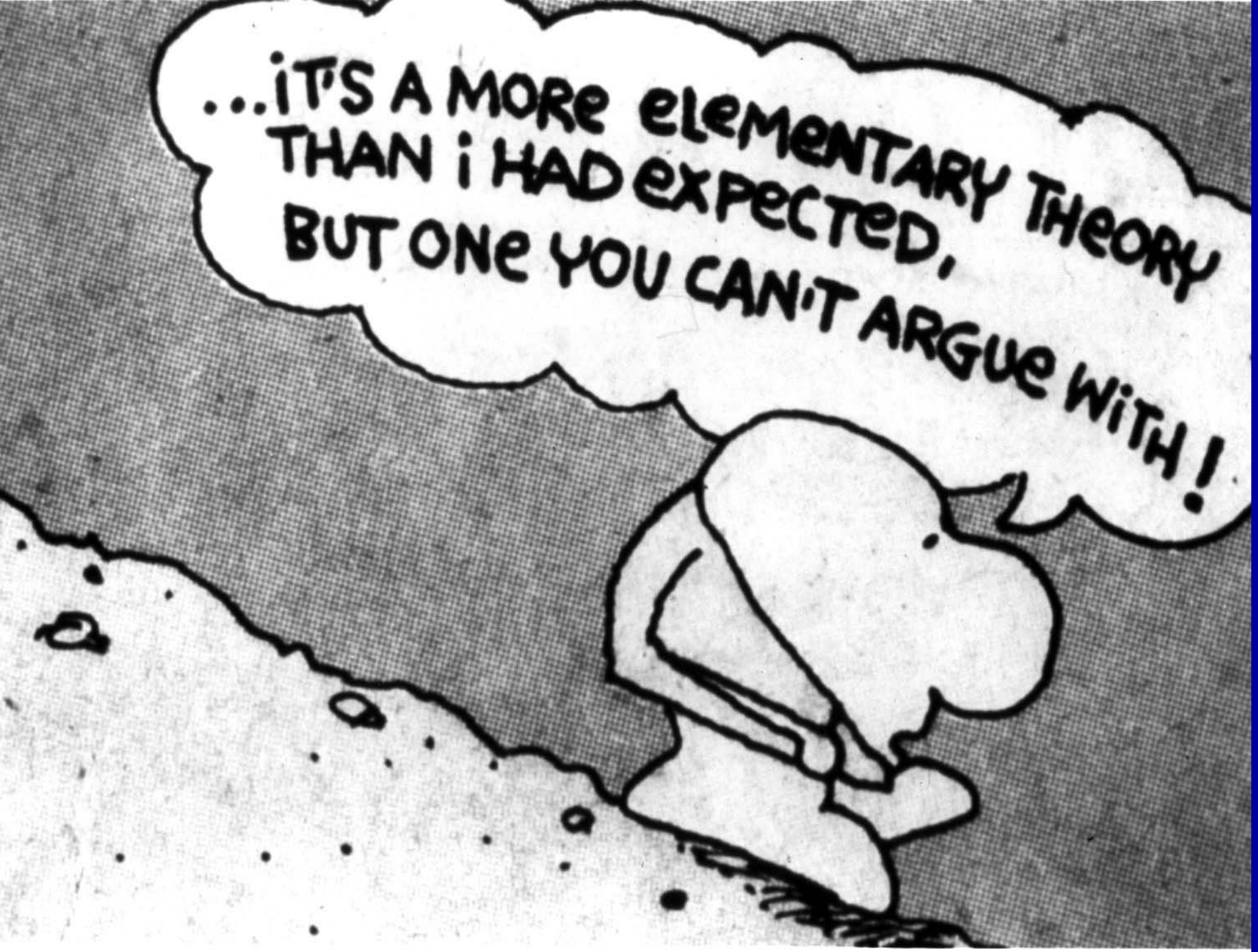
What is the meaning of life?


TELL ME OLD WISE ONE...
WHAT IS THE MEANING
OF LIFE ?
Z Z


AH YESSS... THE MEANING OF LIFE..
LIFE, MY BOY, IS DOIN' STUFF!!


LIFE IS "DOIN' STUFF"??
...THAT'S IT ???


...AS OPPOSED TO DEATH, WHICH IS NOT DOIN' STUFF !!

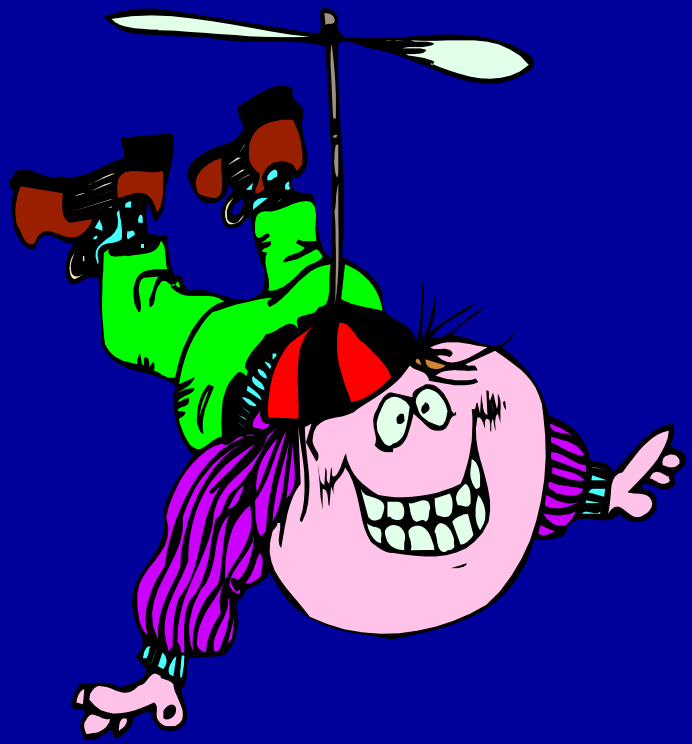

...IT'S A MORE ELEMENTARY THEORY THAN I HAD EXPECTED, BUT ONE YOU CAN'T ARGUE WITH!

Outcomes Measurement

- Does the health care you give, affect patient health status?
- How do you know?
- How do you distinguish between + and - effects on health status?
- OVERALL, does the patient benefit from the health care they are given?

(From Kind, 1995)

Types of HRQOL Measures



- Profile
  - Generic
  - Targeted
- Preference-based

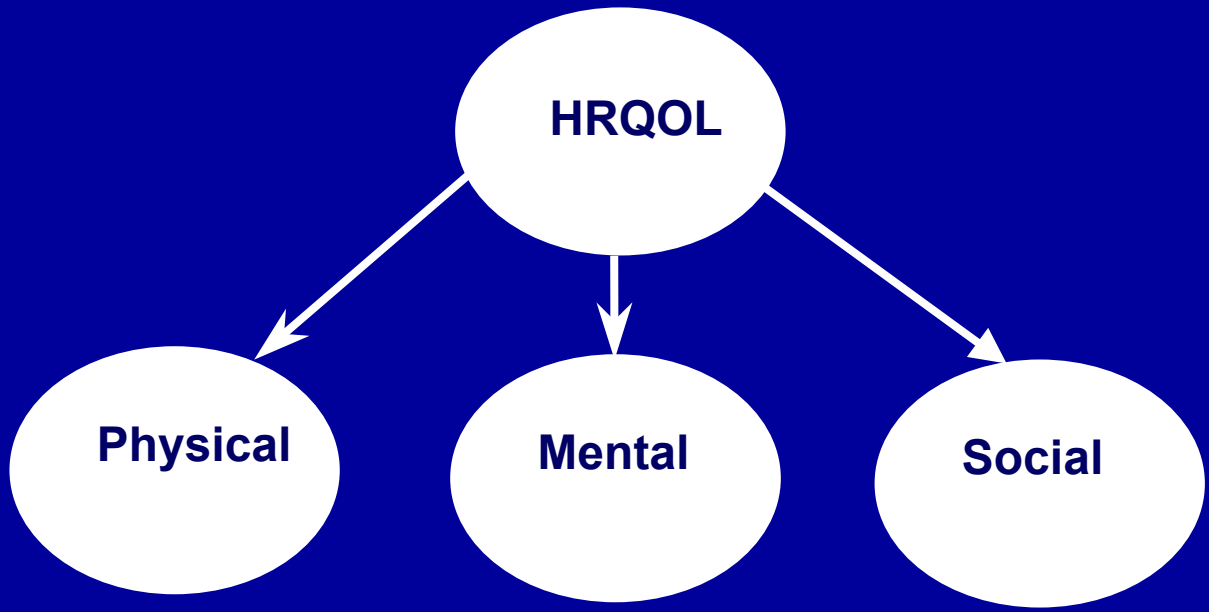
Question 2

Are there generic measures for health-related quality of life?

Health-Related Quality of Life is:

- **What the person can DO (functioning)**
  - **Self-care**
  - **Role**
  - **Social**
- **How the person FEELs (well-being)**
  - **Emotional well-being**
  - **Pain**
  - **Energy**

HRQOL is Multi-dimensional



RAND-36 Scales (Items)

- **Physical functioning (10 items)**
- **Role limitations/physical (4 items)**
- **Role limitations/emotional (3 items)**
- **Social functioning (2 items)**
- **Emotional well-being (5 items)**
- **Energy/fatigue (4 items)**
- **Pain (2 items)**
- **General health perceptions (5 items)**

Physical Functioning Item

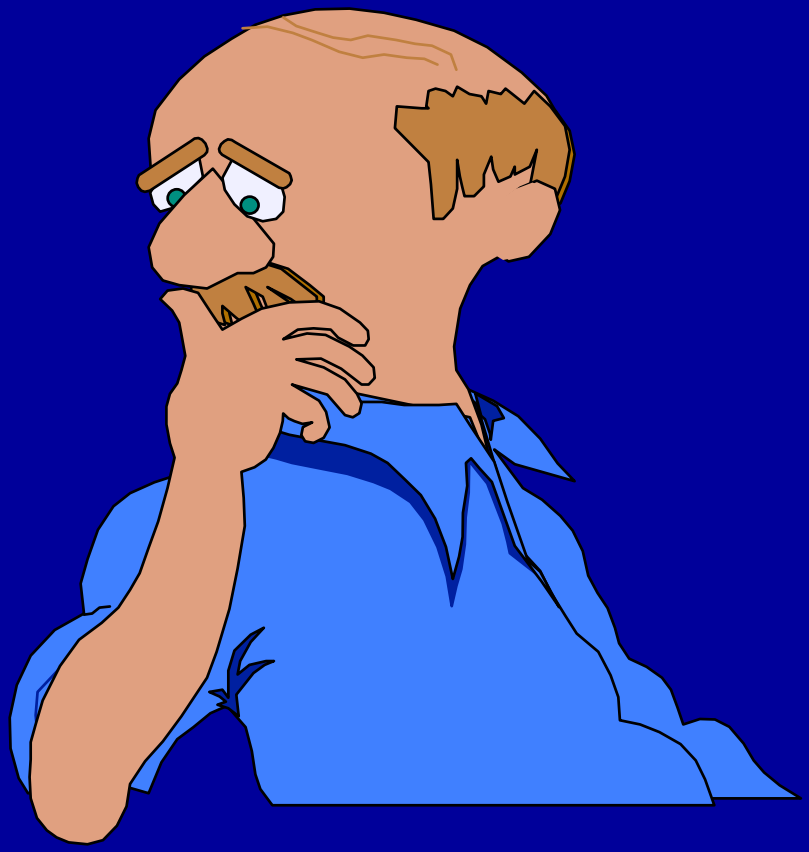
Does your health now limit you in bathing or dressing yourself?

Yes, limited a lot

Yes, limited a little

No, not limited at all

Emotional Well-Being Item



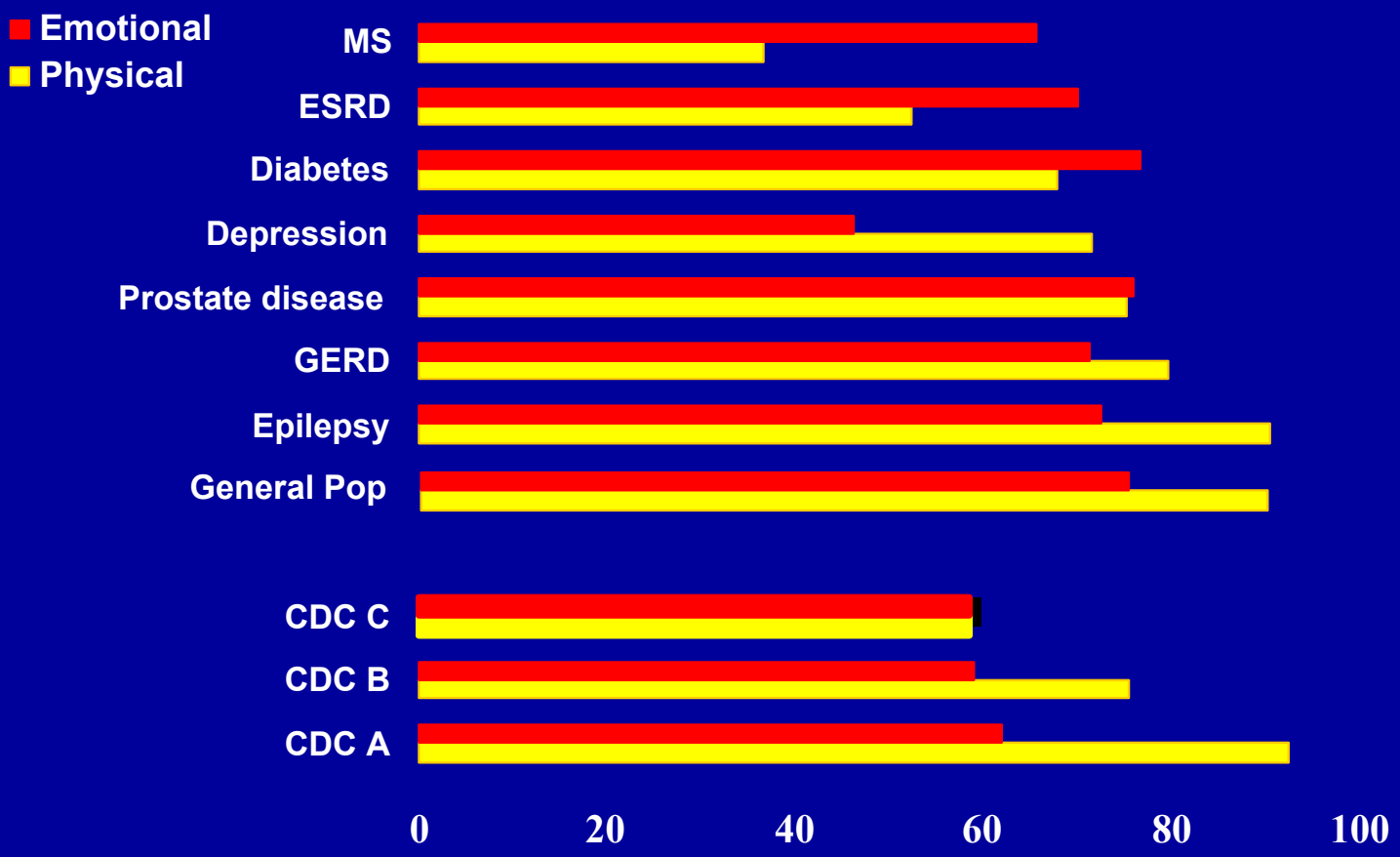
How much of the time during the past 4 weeks have you been a very nervous person?

None of the time; A little of the time; Some of the time; A good bit of the time; Most of the time; All of the time

Scoring RAND-36 Scales

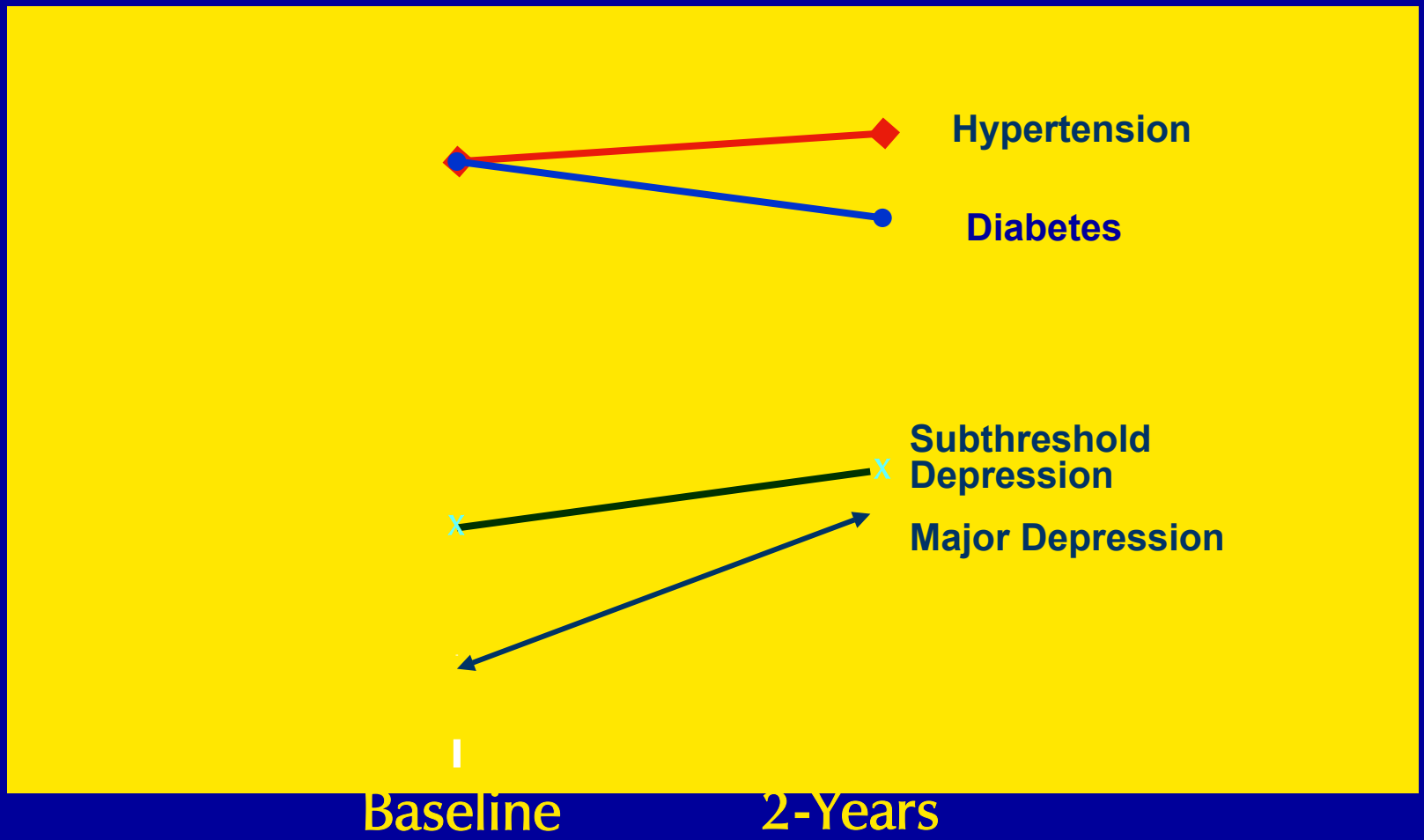
- **Average or sum all items in the same scale.**
- **Transform raw average or sum to 0-100 possible range (linear transformation)**
  - **(raw score – minimum)* 100/(max – min)**

HRQOL of HIV Infected Adults



Hays, et al. (2000), American Journal of Medicine

Course of Emotional Well-being Over 2-years for Patients in the MOS General Medical Sector

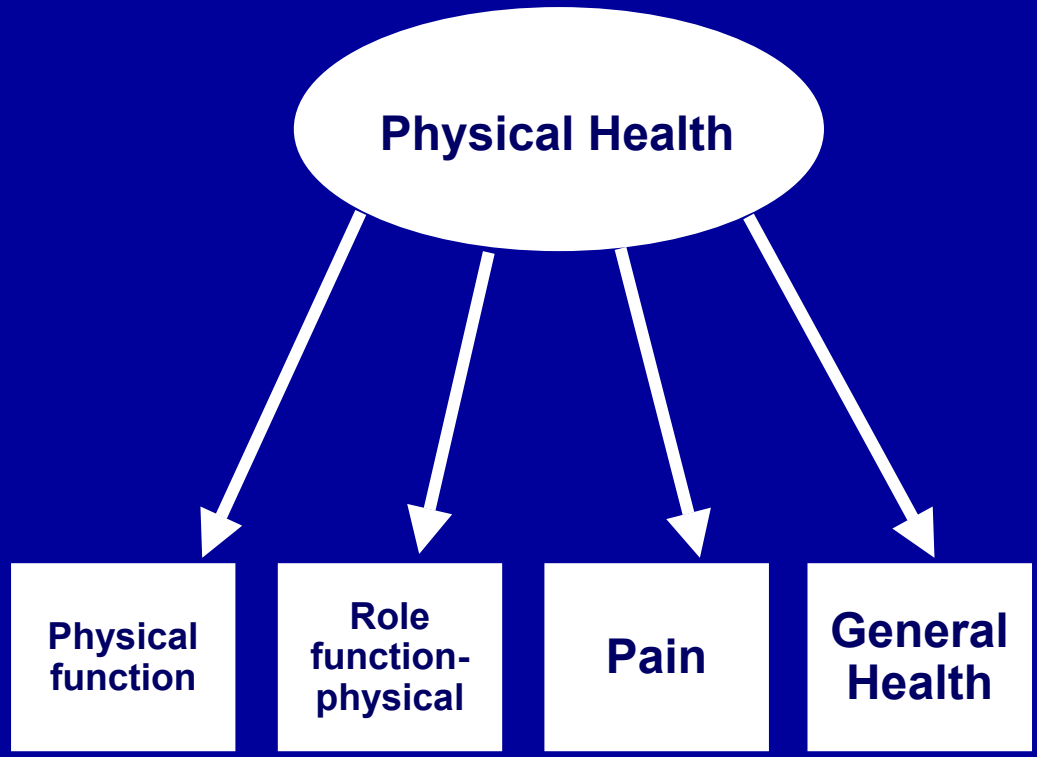


. (1995). Functioning and well-being outcomes of patients with depression compared to chronic medical illnesses. Archives of General Psychiatry, 52, 11-19.

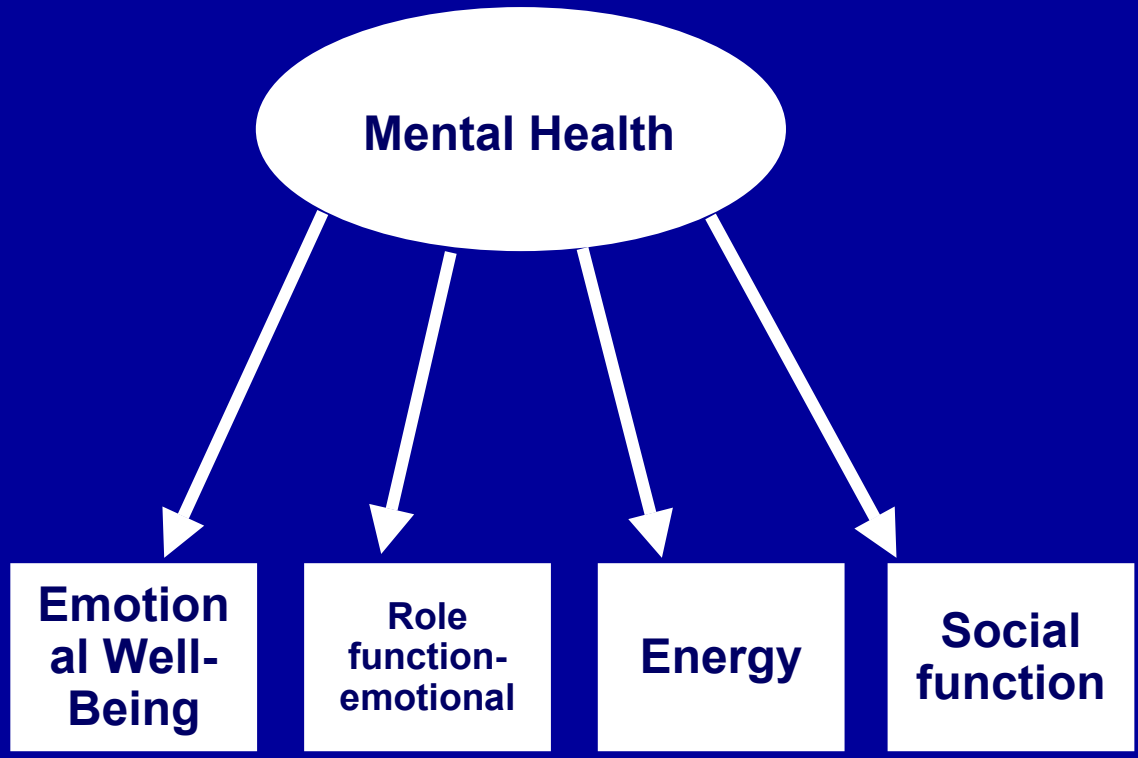
Two Underlying RAND-36 Dimensions

- **Hays, R.D., and Stewart, A.L. (1990). The structure of self-reported health in chronic disease patients. Psychological Assessment, 2, 22-30.**
- **Hays, R. D., Marshall, G. N. et al. (1994). Four-year cross-lagged associations between physical and mental health in the Medical Outcomes Study. Journal of Clinical Psychology, 62, 441-449.**

Indicators of Physical Health



Indicators of Mental Health

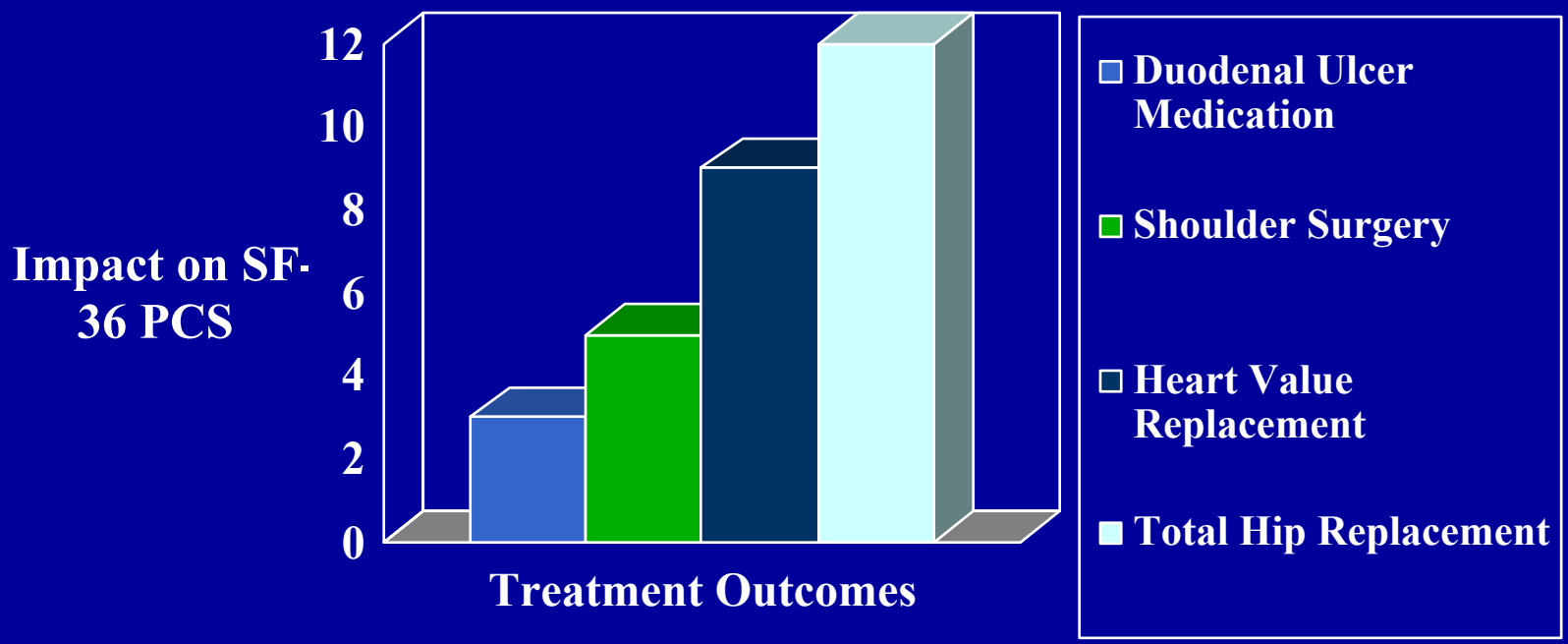


RAND-36 Summary Scores

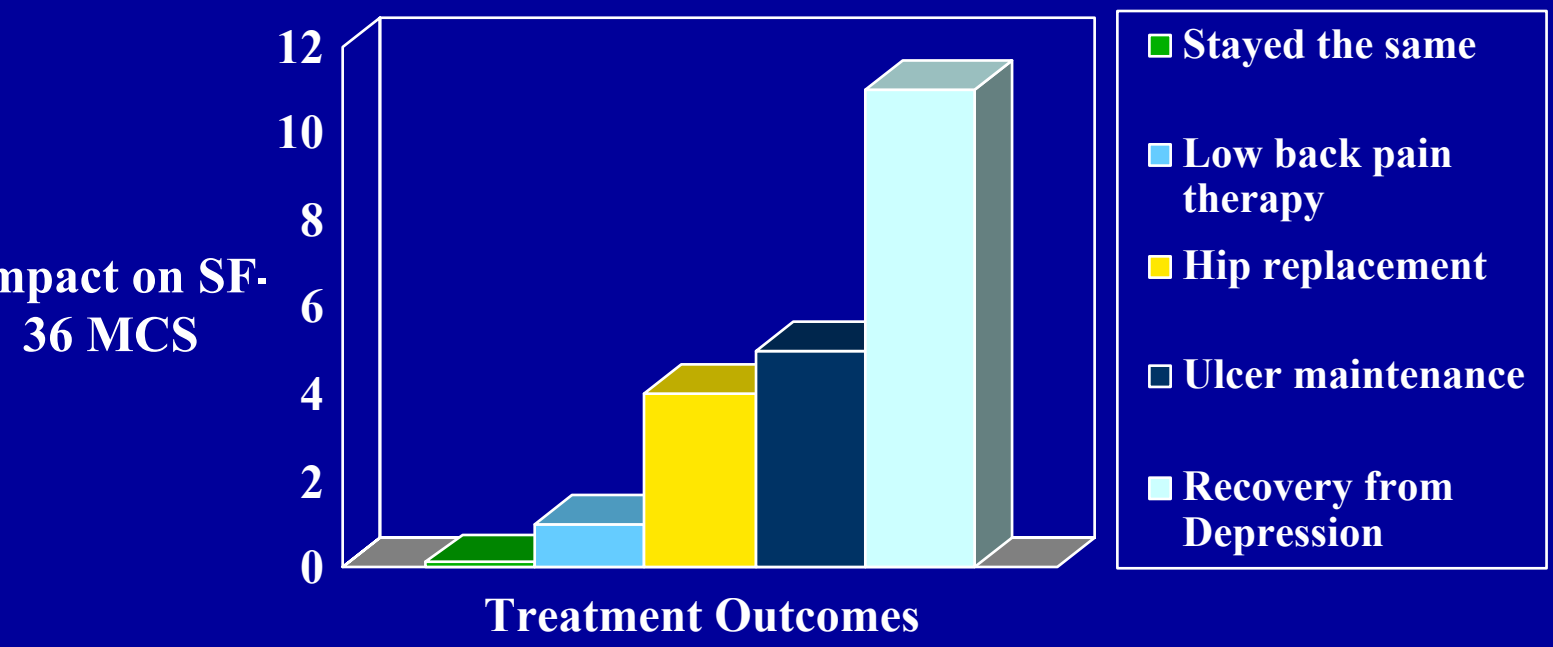
- Physical Health Composite
  - Physical functioning, role—physical, pain, general health perceptions
- Mental Health Composite
  - Emotional well-being, role—emotional, social functioning, energy/fatigue
- Intercorrelation = 0.66; reliability >= 0.91

Hays, R. D., Embury, S. & Chen, H (1998). RAND-36 Health Status Inventory. San Antonio: The Psychological Corporation.

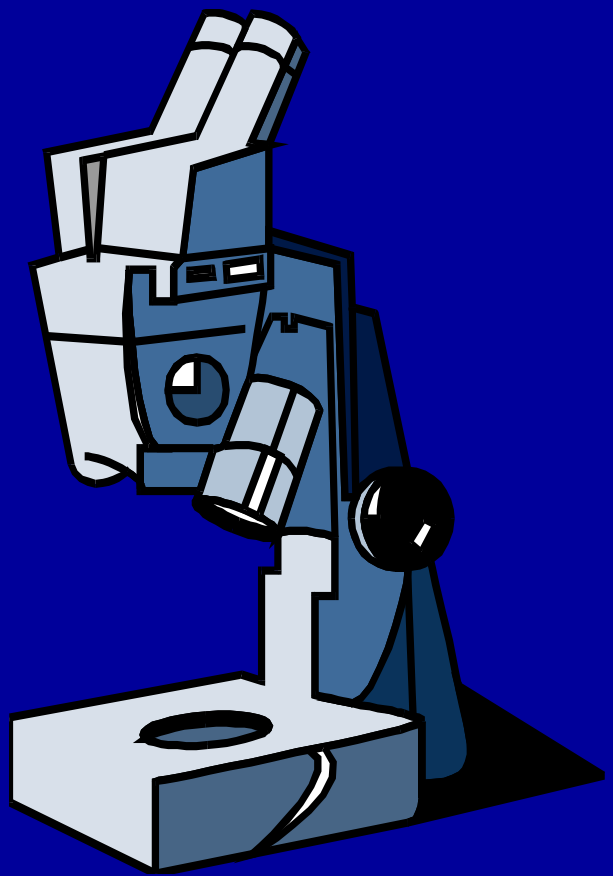
Range of Treatment Impacts on PCS



Range of Treatment Impacts on MCS

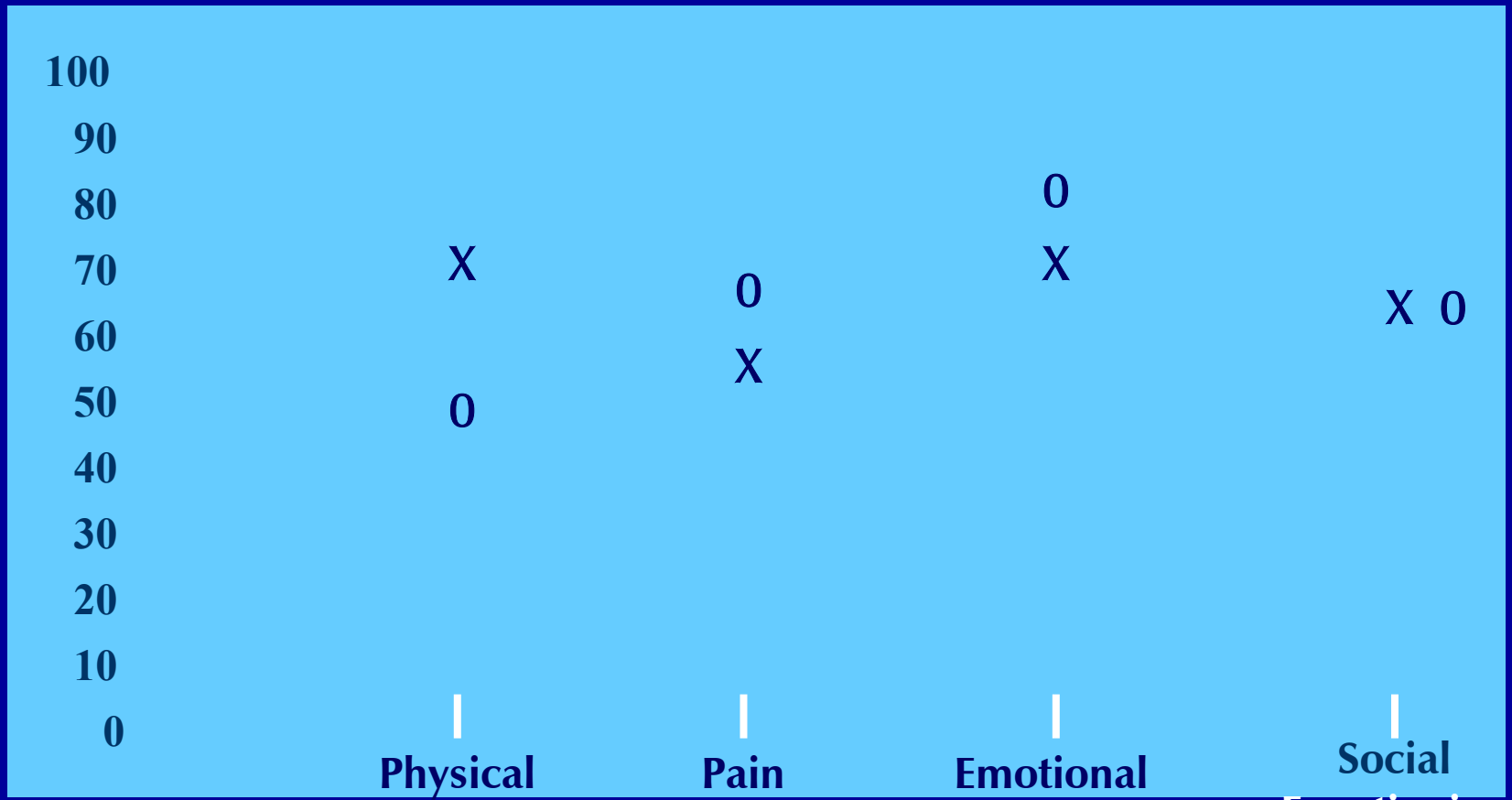


Samsa et al. (1999). Pharmacoeconomics



- MCID for SF-36 is “typically in the range of 3 to 5 points” (p. 149).
- .09->0.28 ES

Limitation of RAND-36: Is New Treatment (X) Better Than Standard Care (O)?



Summary of RAND-36

- Generic profile measure
- Includes eight subscales
  - 4 represent physical health
  - 4 represent well-being or mental health
- Available in many languages
- Remains the most commonly used measure in the world

Utility Approaches

HUI

QWB

EQ-5D

Traditional

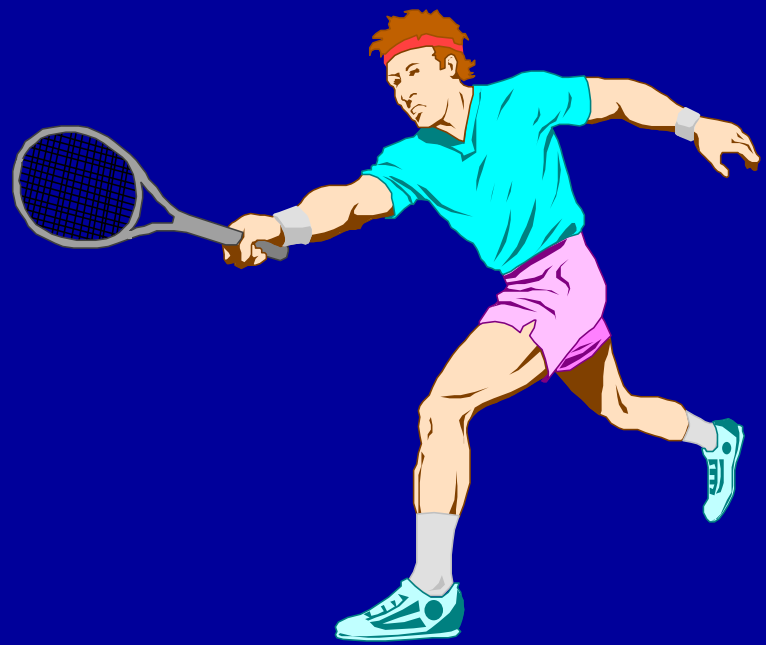
- Life Expectancy
- Infant Mortality
- Disability Days

Survival Analysis

- Alive 1.0
- Dead 0.0

Problem with Survival Analysis

- Tennis player 1.0
- Man in coma 1.0



Quality of Well-being Scale

- Currently two versions
  - Interviewer
  - Self-Report
- Takes about 10 minuets
- Automated scoring, low cost
- About 250 published papers describe use

QWB Components

- Functional Scales
  - Mobility (MOB)
  - Physical Activity (PAC)
  - Social Activity (SAC)
- Symptom/Problem Complexes (CPX)

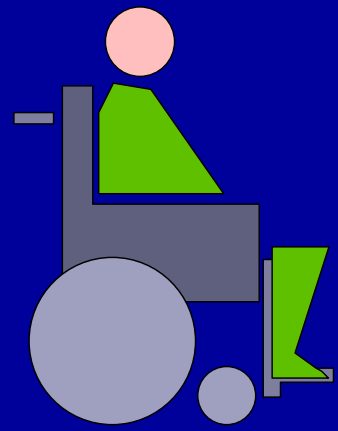
Purpose of Quality Adjusted Survival Analysis

- To summarize life expectancy with adjustments for quality of life

Mobility Scale

- No limitations in travel
- Did not drive or use public transportation
- In house
- In hospital
- In Special Care Unit

Physical Activity Scale



- Walked without physical problems
- Walked with limitations
- Moved own wheelchair without help
- Confined to bed or chair

Social Activity Scale

- Did work, school or housework and other activities
- Did work, school or housework, but limited in other activities
- Limited in amount or kind of work, school, or housework
- Performed self-care, but not work, school, or housework
- Had help with self care

Symptoms or Problems (selected)

- coma
- trouble learning, remembering, or thinking clearly
- pain in back or neck
- sick or upset stomach
- coughing wheezingof breath
- spells of feeling upset, depressed or of crying
- overweight
- runny nose
- problems with sexual interest or performance

Quality-Adjusted Life Year

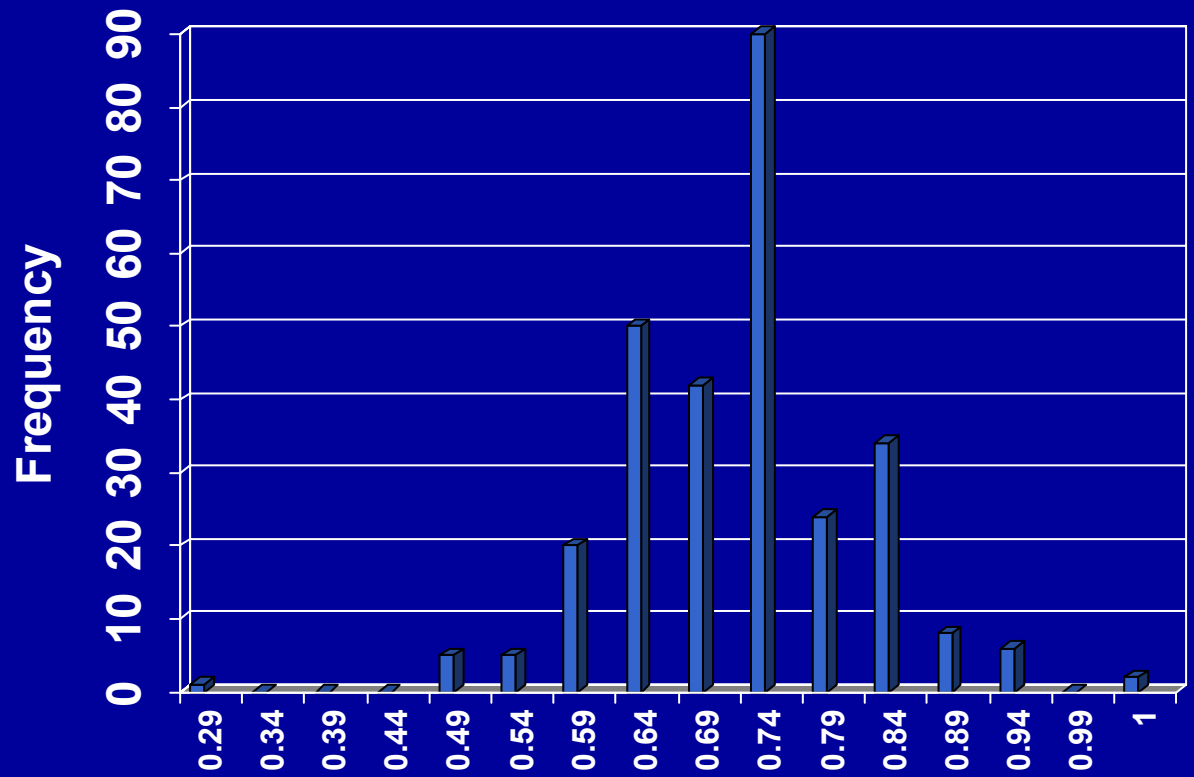
- Combines morbidity and mortality into a single index
- Represents life expectancy with adjustments for quality of life
- Is defined as a year of life free of all disabilities and symptoms

Example Case: 68 year old COPD patient

Description

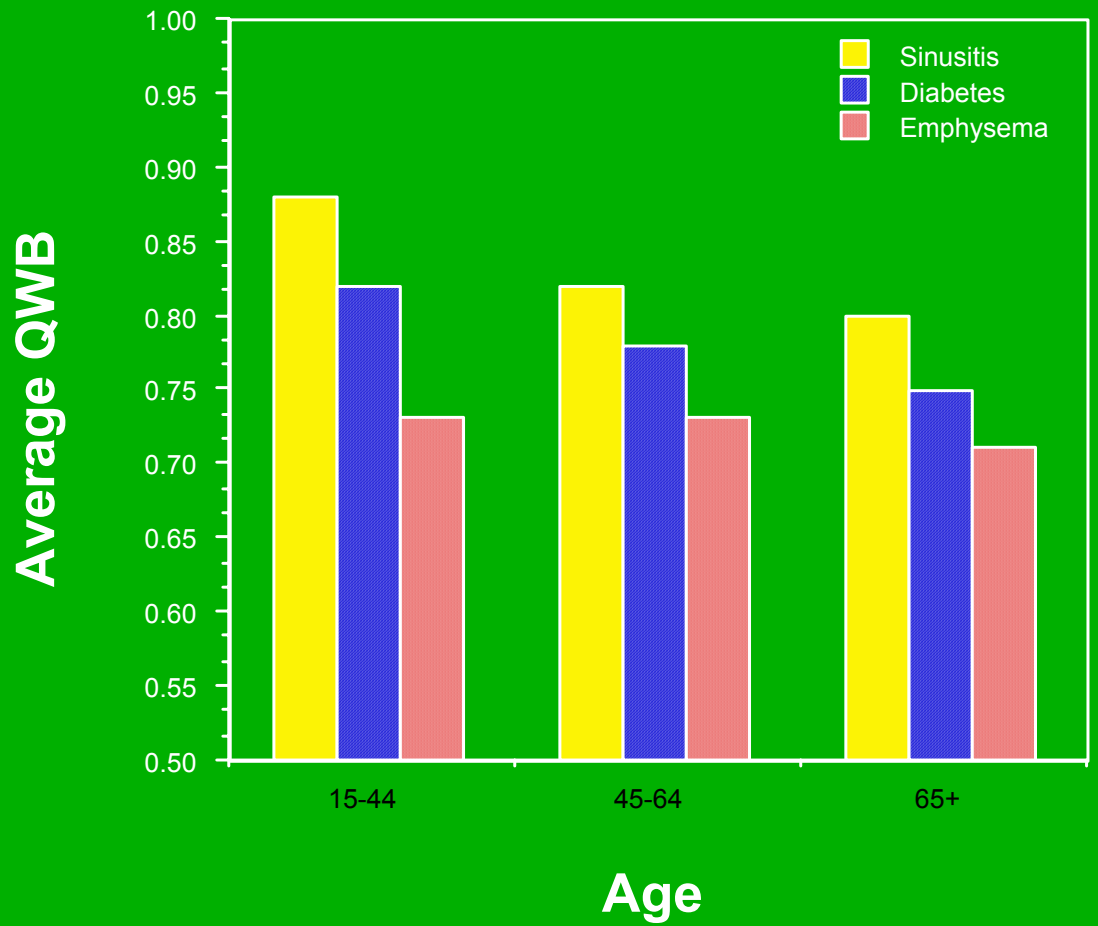
- Shortness of breath
- Drove Car
- In Bed or Chair for Most of Day
- Performed No Major Role Activity, but did perform self-care
- Weight
- Peer Rating equals .605
- For each year in this state, the patient loses 1 - .605 = .395 well years

QWB-SA Distribution (Andresen 1998, N=301)

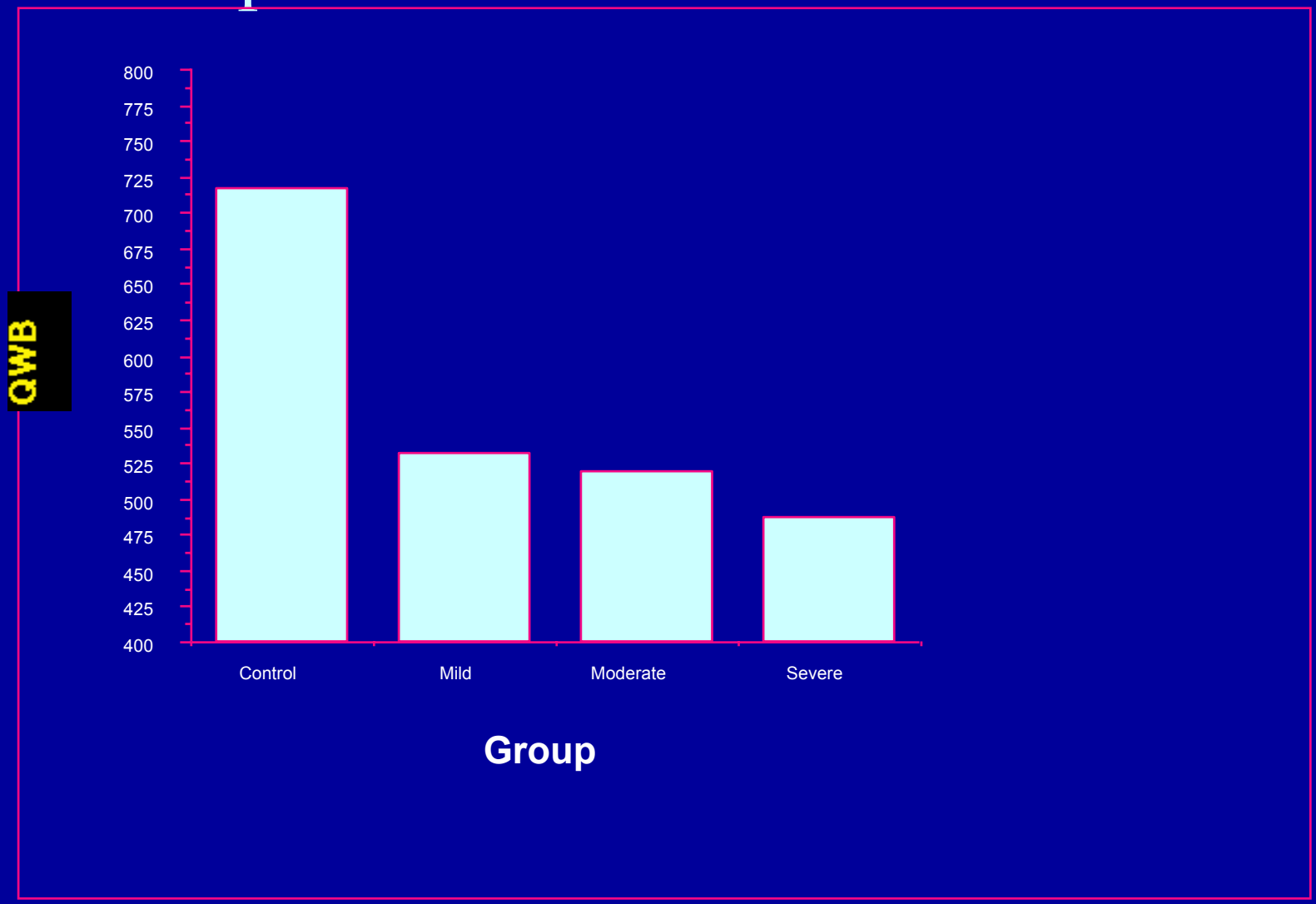


Sinus Disease and Diabetes in the General Population

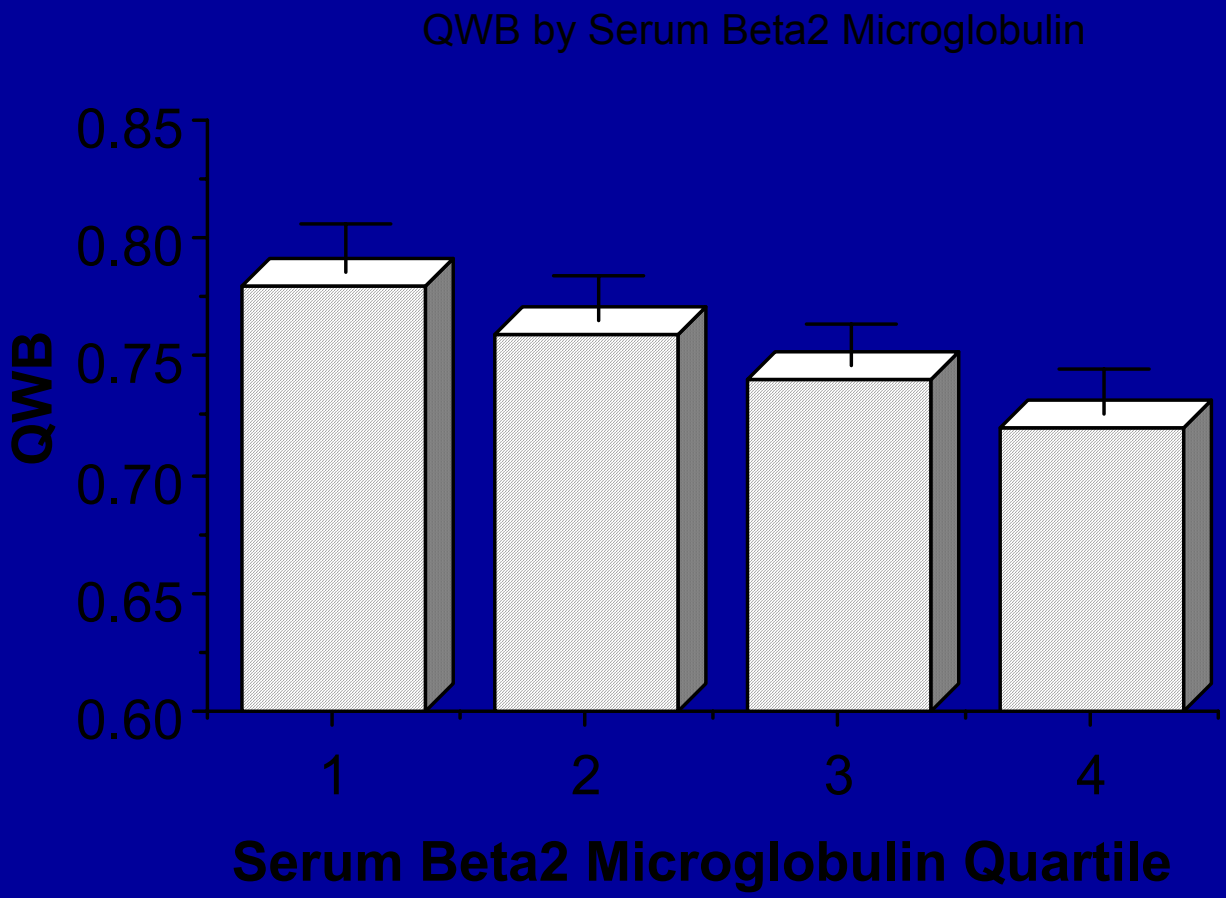
Source: Erickson, 1980 NHIS, Preliminary



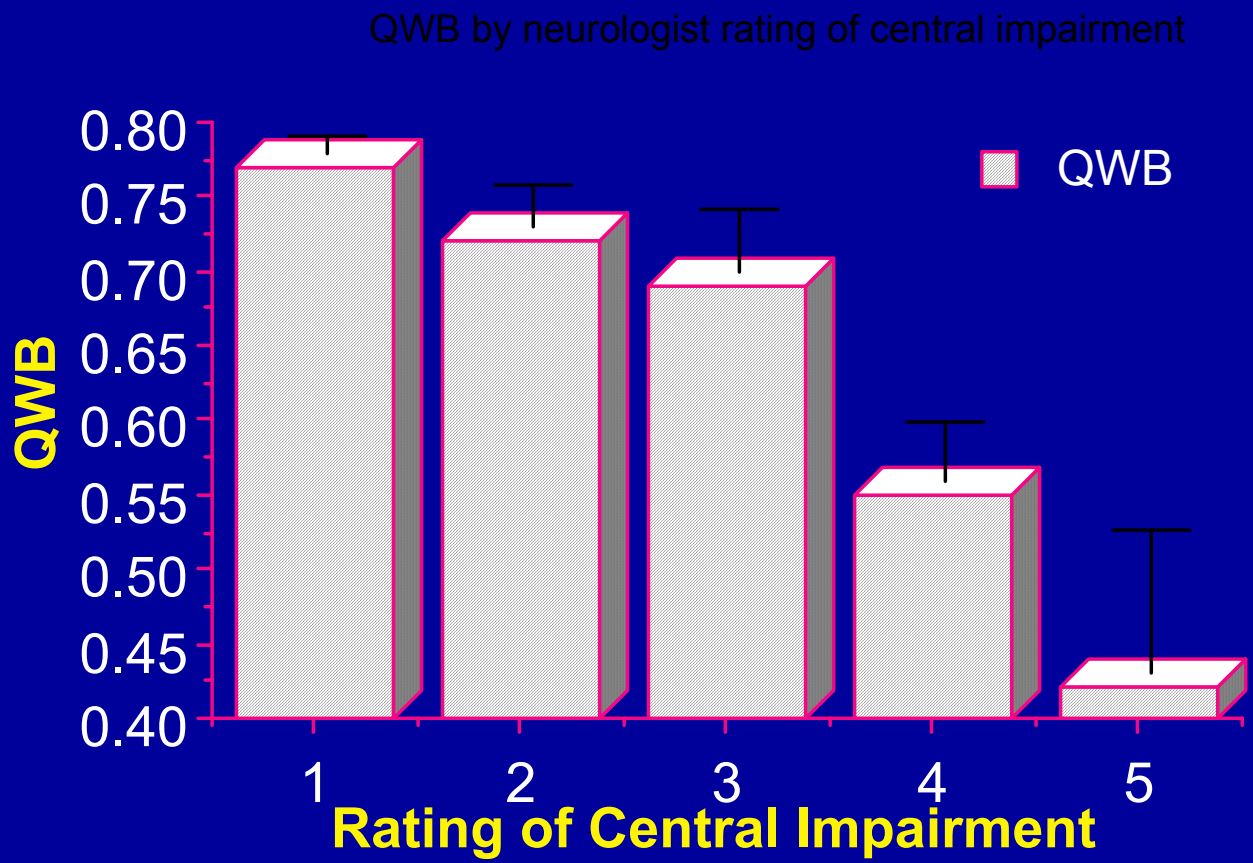
QWB by Level of Cognitive Impairment in Alzheimer's



QWB and Serum Beta 2 Microglobulin in HIV

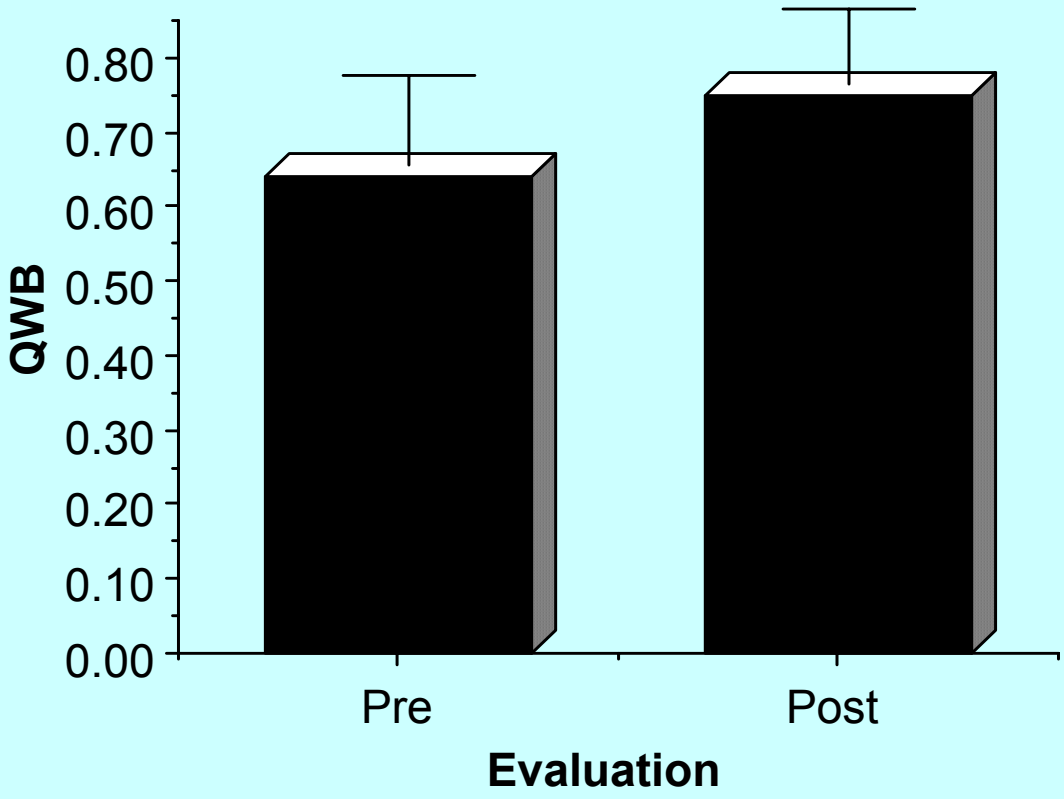


QWB and Neurological Evaluation in HIV



QWB and Survival in HIV

QWB Before and After Ciprofloxacin Treatment for Exacerbations of CF (Orenstein et al, 1990)



QWB by SAPS Patient Groups and Controls

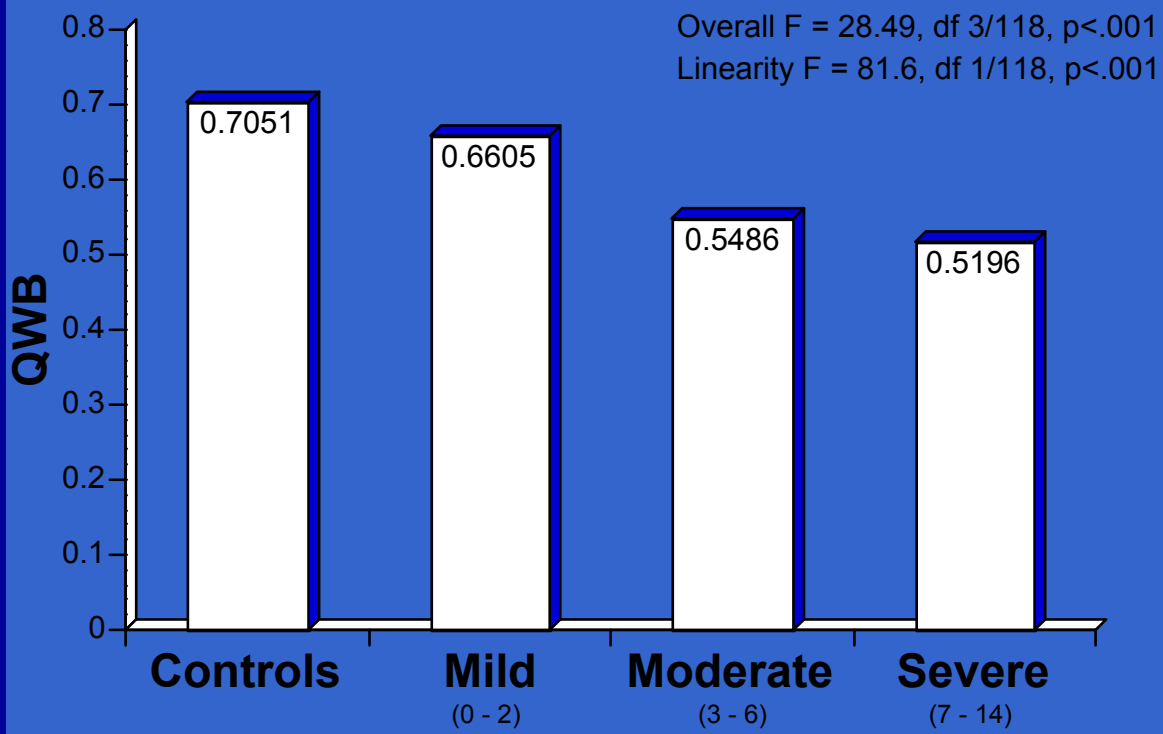
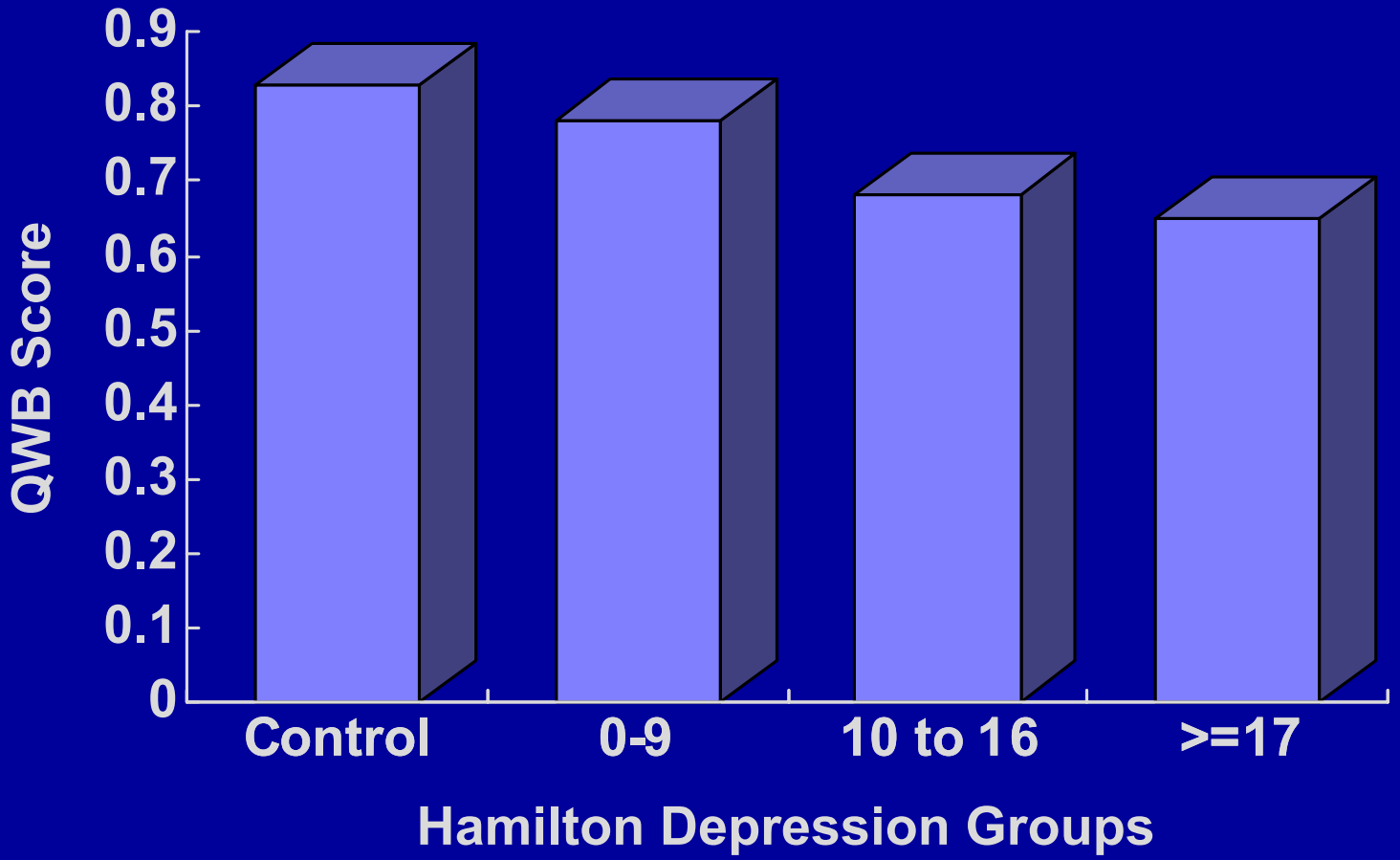
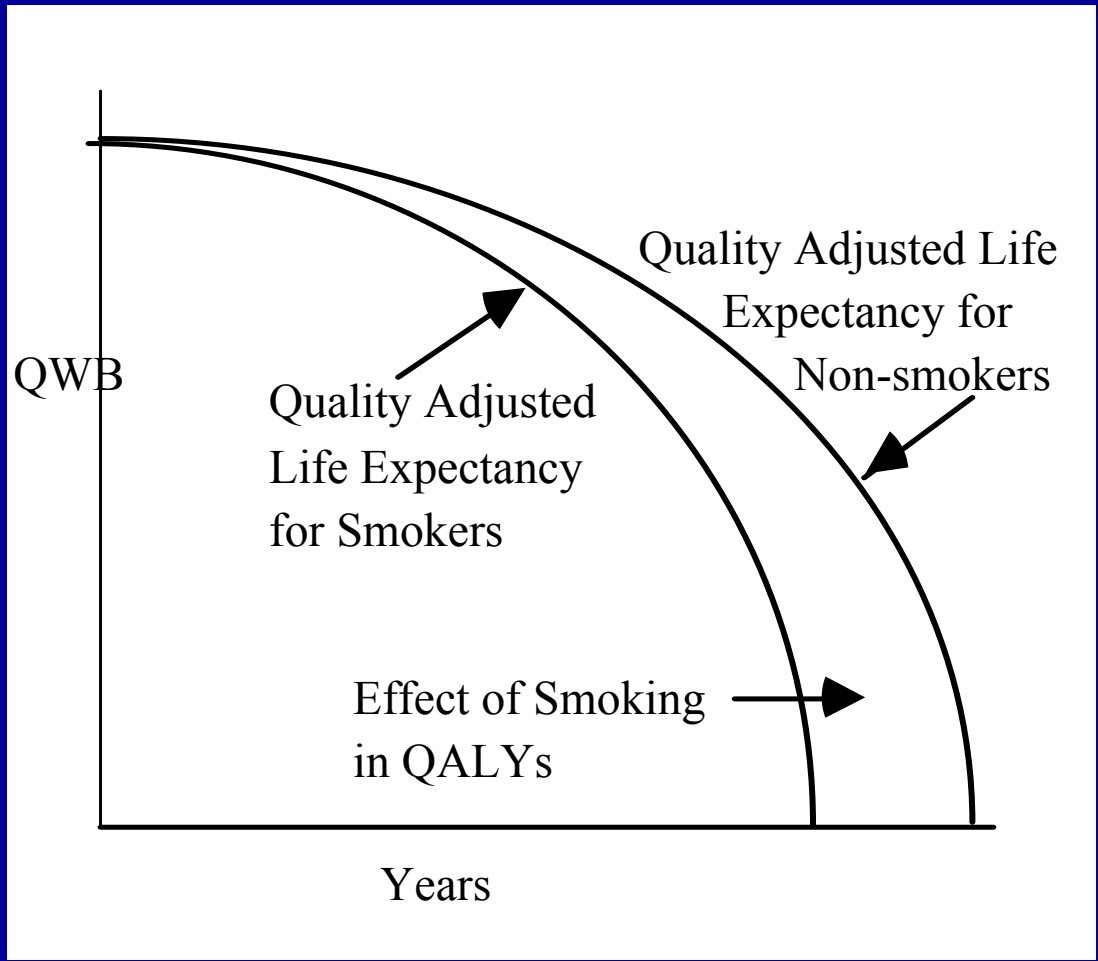


Figure 1

QWB by Hamilton Depression (from Rubin et al 1994)



Estimating treatment effects



Summary

- QWB and SF-36 have some common roots
- Correlations between QWB and some SF-36 components are substantial
- QWB now can be self-administered
- QWB can be used to estimate QALYs for policy analysis
- Several theoretical and technical issues must be resolved in future studies

Comparison

- SF-36 can not be used for cost/effectiveness analysis
- QWB does not offer a profile of clinical outcomes

What if you used the SF-36, but need utility scores for cost/utility analysis?

- The Fryback method is based on the regression of SF-36 components upon the Quality of Well-Being (QWB) scores.
- The Nichol method uses a similar methodology to estimate Health Utility Index (HUI) scores.
- The Brazier method uses original utility ratings to estimate health state evaluations for 1,800 states that could be derived from the SF-36

Fryback et al. Prediction of QWB from SF-36

- 56.9% of the observed QWB variance; 49.5% on on cross-validation

QWB~ = **0.59196**

+ (PF * **0.0012588**)

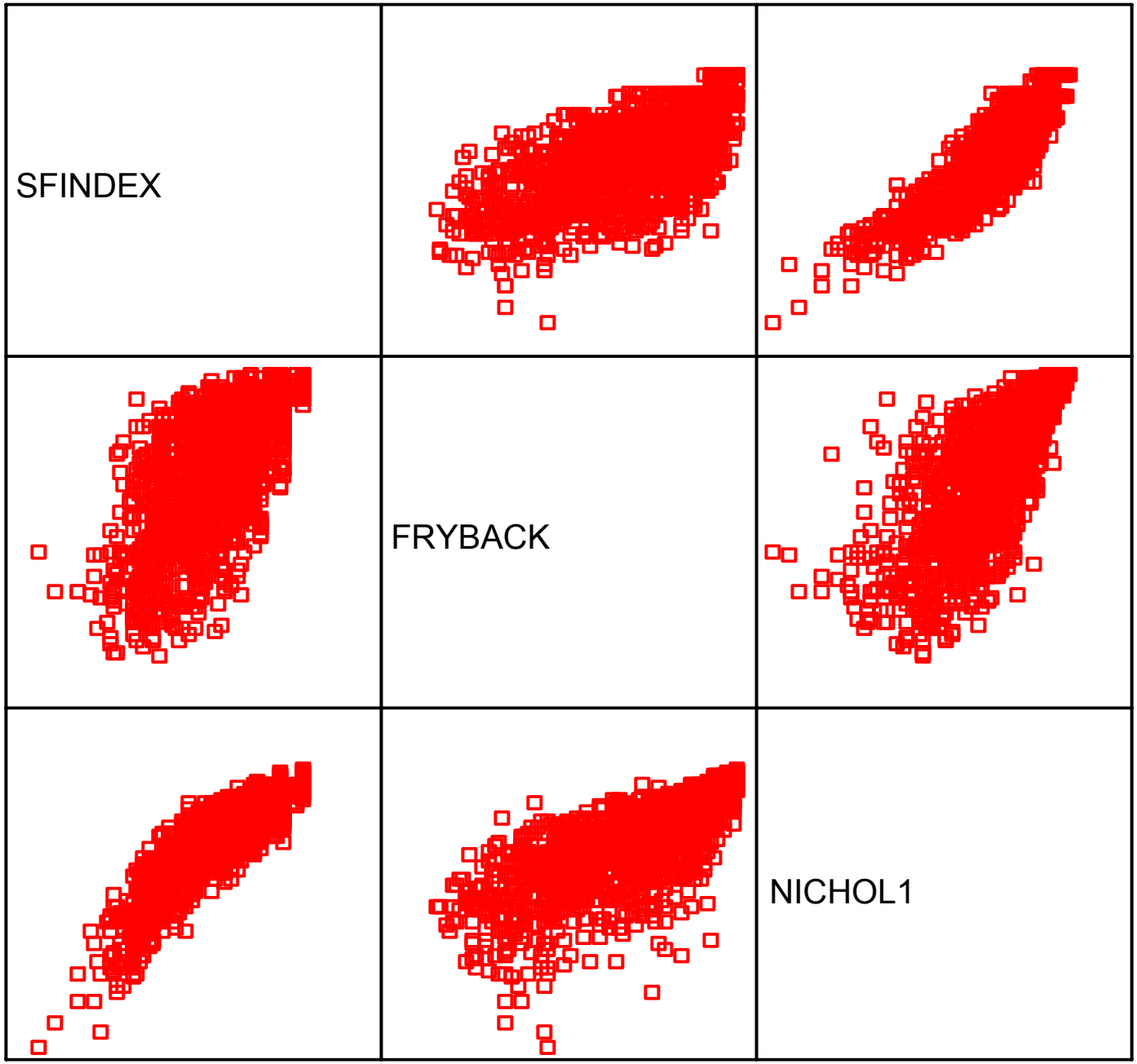
- (EWB * **0.0011709**)

- (BP * **0.0014261**)

+ (RP *x* GH * **0.00000705**)

+ (PF *x* BP * **0.00001140**)

+ (BP *x* EWB * **0.00001931**)


SFINDEX
FRYBACK
NICHOL1

Question 4

How can we best use our resources to improve public health?

Level of Economic Analysis

- Macro level--informs policy
- Micro level-informs clinical decisions

Example Macro Problem

- Oregon - late 1980s
- Medcaid costs were increasing 25% per year
- Medicaid coped with the problem by changing eligibility threshold
- Number of people covered reduced to 200,000 among 600,000 eligible
- Proposed rationing services rather than people
- Goal was to increase number covered

Macro Level Decision

- Fixed level of resources
- Potentially infinite demand
- Need to make effective/efficient use of resources
- Set priorities-make choices

Micro Level Decision

- I am 82 years old
- I feel good and my memory is fine
- My doctor says I have >85% stenosis of my carotid arteries
- She wants to operate ASAP
- She says I may die from the surgery
- She also thinks I may die of a stroke
- What should I do?

Overview

- Cost-utility analysis
  - Effectiveness measured as Quality Adjusted Life Years
- Societal Perspective
  - Related medical and nonmedical costs included
- Time Horizons
  - Primary: within trial
  - Secondary: projected 5- and 10-year outcomes

Resources and Valuation

Cost Element	Source
Medicare-covered services	Medicare reimbursements
Study-related drugs	AWP less 15% acquisition + dispensing fee
Travel costs	Federal travel reimbursement per mile
Patient time	Wages for persons ≥ 65 Bureau of Labor Statistics
Caregiver time	Wages for persons ≤ 65 Bureau of Labor Statistics

*Adjusted to 2002 constant dollars (medical component of CPI, July 2002)

Quality-Adjusted Life Years

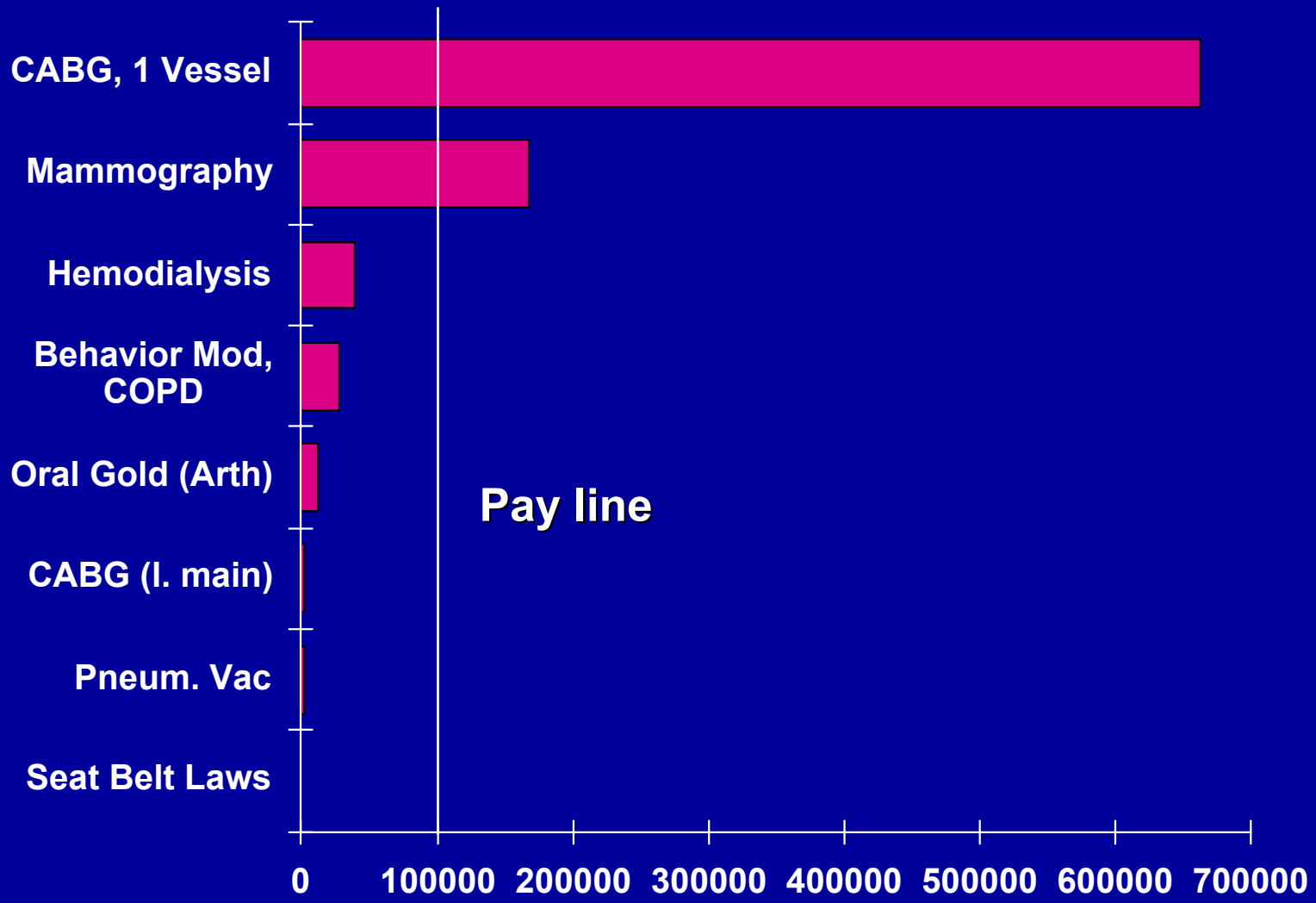
- Survival adjusted for quality of life
  - Range: 0 (death) to 1 (ideal health)

- Quality of life measured as utilities
  - Derived directly from Quality of Well Being scores

Differences Between CBA,CEA, and CUA

Type	**Resources Measured in**	**Outcomes Measured in**
Cost/Benefit	$s	$s
Cost/Effectiveness	$s	Clinical measure (ie mmHg
Cost/Utility	$s	QALYs

Cost/QALY for Selected Interventions



Opportunity Costs; Life years/$1 million, IOM 1999

• Influenza vaccine, persons 65+	7,750
• Smoking cessation	217
• Lovastatin 20 mg men total chol 300	42
• Captopril for hypertension	8
• Pap smear every 3 years	36
• Pap smear every year	<.5

Misconceptions about Cost/utility analysis

- CUA is usually used to justify cutting budgets
- CUA will damage patients
- CUA is about saving money
- CUA is neutral with regard to budget.
- Use of CUA should result in improved population health
- CUA is about saving lives

Question 5

Is there consensus about the which methods should be used?

What has held us back?

- Distractions
  - Disagreements on which measure is best
  - Disagreements on general philosophy of outcome measurement
    - Generic vs disease specific
    - Psychometric vs. utility based
    - Disciplinary differences – statistics, economics, medicine, psychology, anthropology….

We do agree on some of the core issues

- Most measures can be traced back to Sullivan (1966)
  - Sullivan rarely cited
- Content of items is remarkably similar
- Most measures combine measures of life length and life quality
- Most quality of life measures are hybrid health status/utility measures
  - Health states and health weights (Erickson)

John Ware

- Think of different approaches as brand names of products designed to measure the same underlying construct… health

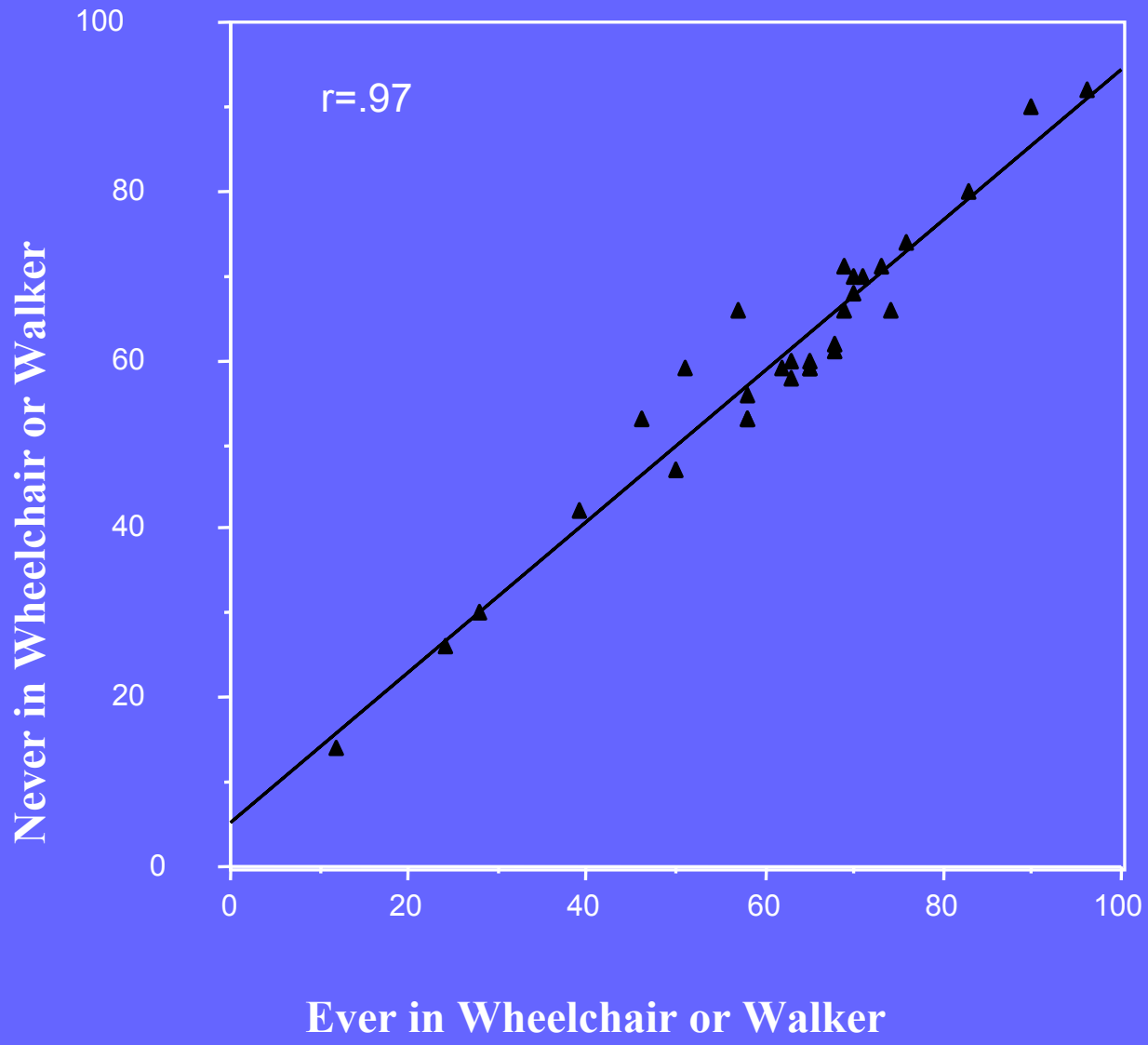
Major Distracter 1: Preference and Utility Assessment

- Standard Gamble
- Time Trade-off
- Rating Scales
- Think scoring systems

Major Distracter 2: Response Shift

- Preferences of patients and non-patients differ
- As a result, preferences weights have no meaning
- But, is this supported by evidence?

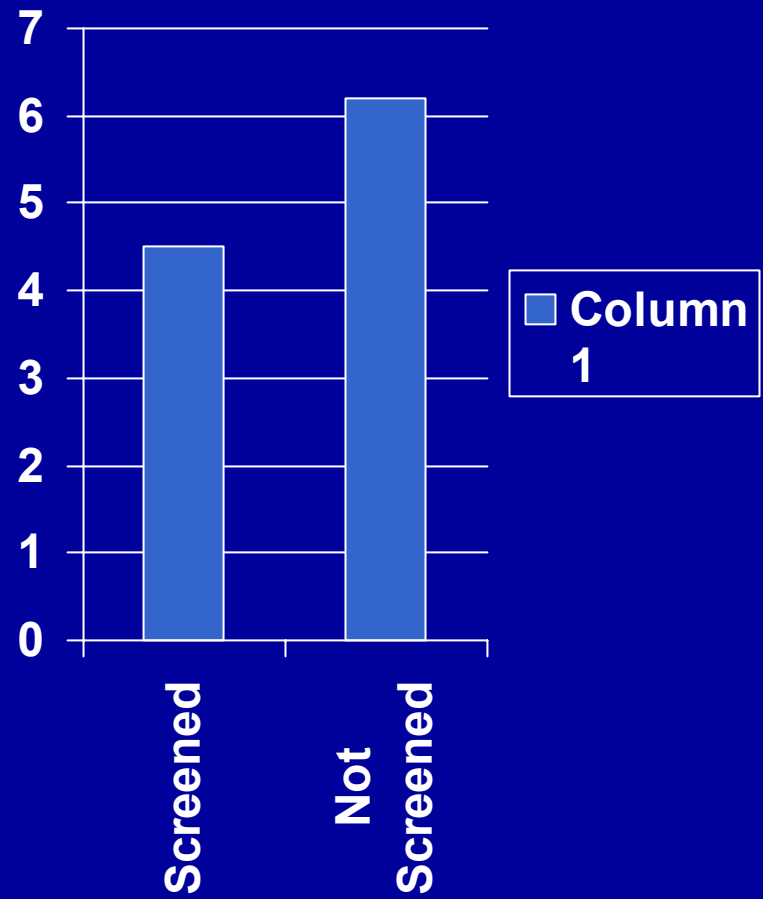
Comparison between ever and never in wheelchair or walker for 31 items:
Data from Oregon Health Services Commission



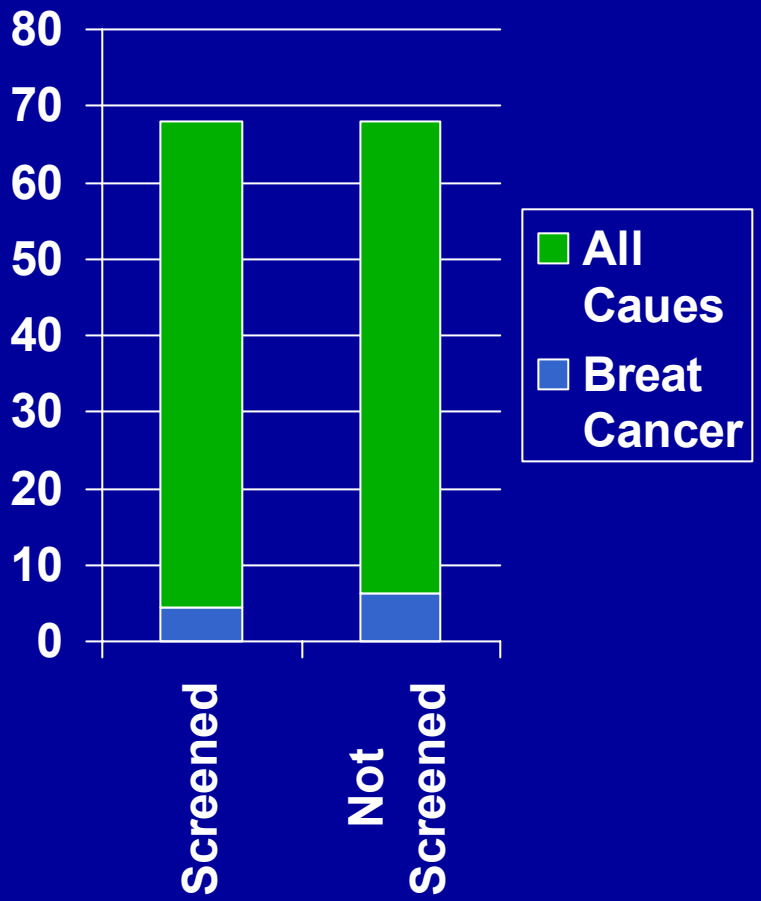
Major distracter 3: The total mortality problem

Cancer mortality in the Health Insurance Plan of New York

- 60,000 women assigned to mammography or usual care
- After 10 years 147 deaths in the mammography group and 192 deaths in usual care group
- 23% reduction in cancer deaths



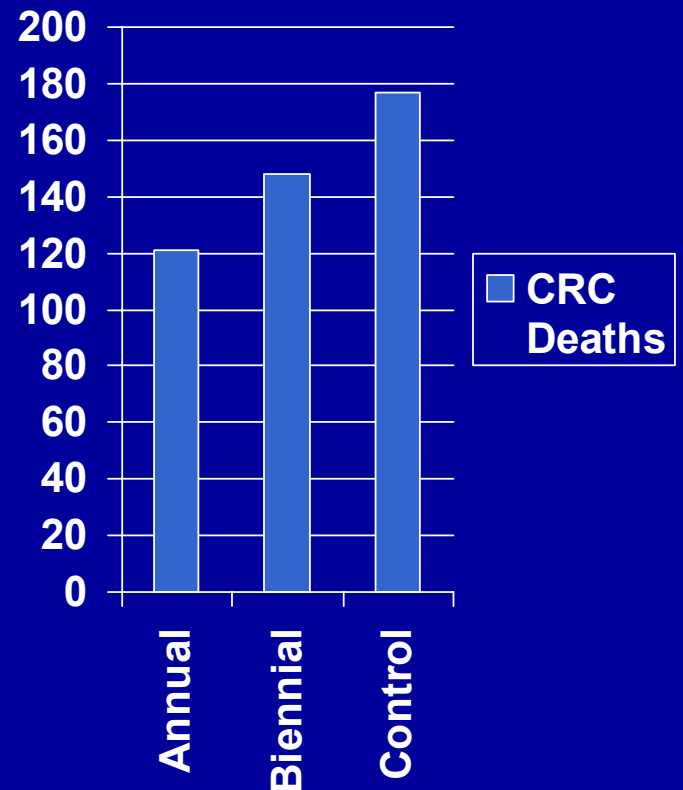
Cancer mortality in the Health Insurance Plan of New York



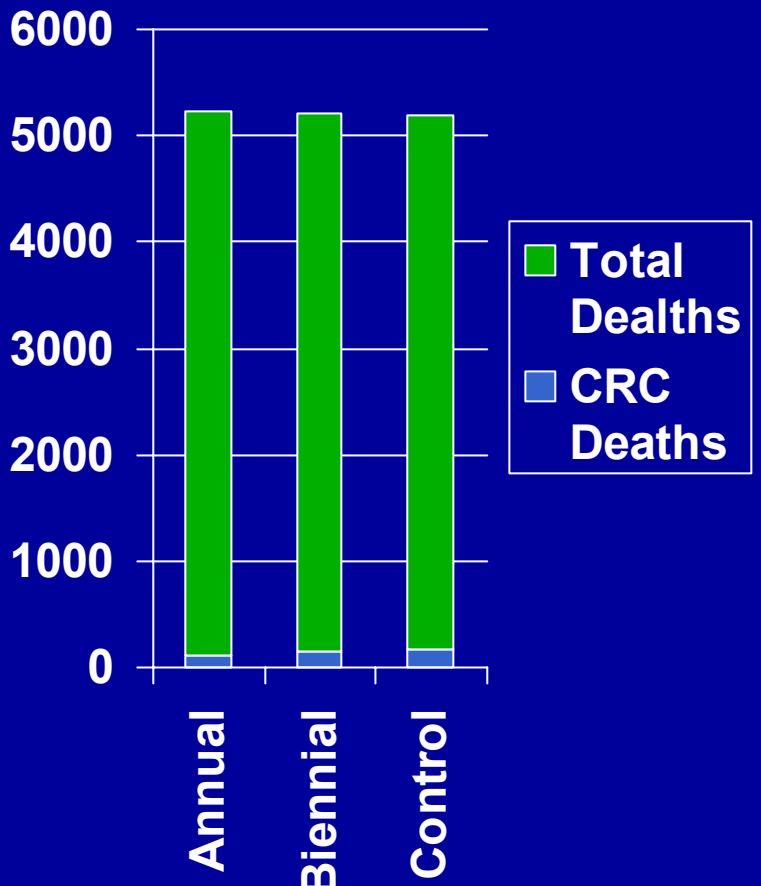
- Lower portion shows cancer deaths, upper shows non cancer deaths
- No difference is survival between screened and unscreened women

Minnesota Colon and Rectal Cancer Screening Study

- Headline "Screening reduces cancer deaths by 32%"
- Over 45,000 participants
- CRC deaths, 121 for annual screening, 148 for biennial, and 177 for control



Minnesota Colon and Rectal Cancer Screening Study



- No differences in total mortality (5236,5213,5186)
- Absolute risk of death .33 for all groups
- Absolute risk of CRC death: Annual .007, Biennial, .009, Control, .011

Summary

- Ziggy– life is about doing stuff
- SF-36 offers a well validated profile of health outcomes
- Utility measures can be used for cost/utility analysis
- Cost/utility analysis will become increasingly important it RCTs